DOCUMENT RESUME

ED 119 075 CG 010 384

AUTHOR Bonoma, Thomas V.
TITLE Social Psychology and Social Evaluation.
PUB DATE [74]
NOTE 31p.; Paper presented at the Annual Convention of the American Psychological Association (83rd, Chicago, Illinois, August 30 to September 2, 1975); Not available in hard copy due to marginal legibility of original document

EDRS PRICE MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS Conceptual Schemes; Conflict Resolution; *Evaluation; Political Influences; *Research Criteria; *Scientific Methodology; *Social Psychology; Systems Concepts; *Theories

ABSTRACT

Recent contentions that the settings and problems of social evaluation research renders this new "subdiscipline" substantively distinct from the remainder of social psychology are critically examined. It is argued that the demonstrable existence of meta-conflicts about the conduct of social evaluation efforts, political sensitivities in the research arena, and the adaptive instability of evaluated programs do not functionally segregate social evaluation from traditional social psychology. Rather, these problems occur homomorphically in both arenas, although they are more often articulated in the former rather than the latter settings. The conclusion of radical social evaluation authors that the scientific method is inapplicable to their research settings in unwarranted: experimental tactics are often degraded by the presence of complex system constraints, but other standard investigatory tactics may be reliably applied toward replicable knowledge. Any segregation of social evaluation from social psychology is dysfunctional, since a relevant social psychological science can be approached only with the development of overarching theory capable of explaining interactive behavior in both settings. Correspondingly, social psychological relevance varies neither with a researcher's investigatory strategy nor problem setting, but directly with the goodness of evolved theory. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED119075

SOCIAL PSYCHOLOGY AND SOCIAL EVALUATION

BEST COPY AVAILABLE

HARD COPY NOT AVAILABLE

Thomas V. Bonoma

Graduate School of Business

University of Pittsburgh

Pittsburgh, PA 15260

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

SOCIAL PSYCHOLOGY AND SOCIAL EVALUATION[1]

(M)en ought to know that in the theatre of human affairs it is only for Gods and angels to be spectators. -- Francis Bacon

Generally, evaluation research may be defined as the conduct of social scientific inquiries, usually in the context of some institution, corporation or agency, where the investigatory purpose is a functional assessment of some unit subsystem. Social evaluation, public (i.e., tax-based) sector evaluation, correspondingly focuses on some program (D. Cook, 1966; Grobman, 1970; Moores, 1973; Taylor, 1973), policy (Evans, 1972; Weiss, 1972, 1973; Wozniak, 1973) or service provided by a "social" agency (DuBois & Mayo, 1970); the functional assessment sought is an index addressing the efficiency or practicality of service delivery or program operation. In both cases, "the essence of evaluation is attribution" (Evans, 1972, p. 634), where attribution is understood as a trained observer's scientifically guided judgment about program worth (cf. Scriven, 1967).

Due to a sharpening accountability focus in the public sector, the last 5 years have seen an increasing demand (cf. the APA Monitor) for social scientists competent to provide empirical indices of program worth. Correspondingly, and prodded by the dual goad of an oversaturated academic market as well as strong demands for research relevance (Silverman, 1971), an increasing number of social psychologists are entering the social evaluation arena either as full-time practitioners or academically-based "social relevance" researchers. The net effect has been the quite recent appearance of a significant (e.g., the 1973, 1974 APA Programs) body of literature treating the "proper" conduct of social evaluation endeavors.[2]

It has currently become fashionable in the social evaluation literature to emphasize the uniqueness of these undertakings as compared to traditional[3] social psychological inquiry. Structurally, at least Argyris (1970),

Hornstein et al. (1971) and Krause & Howard (in press) have coined neologisms (e.g., "social intervention") designed to encompass social evaluation concerns as well as delimit them from the remainder of social psychology. Specialized journals, and even department titles complement this differentiation.

The functional differentiation is not limited to mere nomological exercises, however. Rather, when one invents a new name for something, one must then take special pains to demonstrate the true uniqueness of the creation. Just such a functional segregation of social evaluation from social psychology has been attempted by a number of recent authors. These assert that the problems and settings of social evaluation are by nature different from traditional pursuits (e.g., Guttentag, 1973; Koen, 1975; Krause & Howard, in press; Sechrest, 1973). Consequently, it is claimed that the adequate performance of social evaluation requires a new specialist (Krause & Howard, in press) versed in novel methodological and conceptual tools "not likely to be found coexisting with very many social scientists today" (Sechrest, 1973; p. 2; cf. Guttentag, 1973; cf. Proshansky, 1974).

This author has no quarrel with those who would separate social evaluation from social psychology, beyond a degree of dismay and a portion of fear. The dismay is recurrent, and generated by any new attempt to shatter our already fractionated discipline still further. The fear, though, is caused by a recognition of pendulum-swing oscillations between proponents of either "side" of the sort which Hull once characterized in another context as "metaphysical and theological controversy" (1935, p. 492). It is true that those with evaluation concerns historically have been adjudged as less "mainstream" than basic researchers (cf. Marx & Hillix, 1963; Samuels, 1973); it is equally

true that, primarily through the efforts of social evaluation pioneers (e.g., Argyris, 1970; Lewin, 1951; Scriven, 1967), "...'applied' concerns' ...are no longer seen as the sordid options of mental cripples" (Koch, 1971, p. 672). However, just such a history makes the hypothesis plausible that current segregatory attempts may represent less de natura differences in problem settings than the zealous slogans of a new and rising sect. We initially examine the alleged differences between the settings and conduct of social evaluation and traditional social psychological research to evaluate this hypothesis.

Some Tests of Parallelism

The literature yields four major issues which purport to demonstrate substantive differences between the settings, problems and conduct of social evaluation and traditional social psychological researches. In overviewing these, we attempt to apply some common-sense tests of homomorphism. If social evaluation writers are correct in their charges of noncomparability with traditional excursions, their pursuits will be seen to involve different (unique) sorts of settings, and encounter different (noncomparable) kinds of difficulties. If, however, the issues cited by evaluation specialists appear to occur also in the traditional domain, even if with differential absolute frequencies or emphasis, we may conclude that the two efforts are homomorphic (i.e., pattern-matched; they "map") and thus not substantively different. The need for "new specialists" and "novel methods" would thereby become somewhat less clear, though a demonstration of homomorphism would not nullify charges that existing methods are insufficient to both research domains.

The first issue raised cites the attributive and advocative nature of social evaluation pursuits as compared to traditional ones (cf. David, 1971). It is claimed that evaluations are specifically designed to operationalize

the articulated values and goals of a program, and are therefore explicitly conducted in order to judgmentally determine to what extent these values have been satisfied. Since traditional research is said to aim toward being "objective" and "value-free," social evaluation should be regarded as a conceptually distinct endeavor from traditional research (e.g., Charlesworth, 1973; Guttentag, 1973; Koen, 1973; Krause & Howard, in press; Weiss, 1973).

This is a simultaneously difficult yet tenuous issue since (a) it disappears at all but the polar cases of the two settings, and since (b) earlier evaluation writers have amply demonstrated their ability to carry out such "attributive" research within the traditional mold (cf. Scriven, 1967). Yet, it ought to be pointed out once more that traditional research is never value-free or objective (e.g., Nagel, 1961; Schlenker, 1974). If we've learned anything from the social psychology of the psychological experiment (e.g., Barber & Silver, 1968), it is that we must opt to make our value biases explicit in research or else suffer the consequences. Secondly, the foci and purposes of traditional research efforts parallel almost exactly the description of social evaluation efforts given above. In fact, the more frequently employed varieties of experimentation (e.g., hypothetico-deductive) have as their formal aim the "evaluation" of some model of real world phenomena (Kaplan, 1964). Finally, in both traditional and social evaluation pursuits, the cycle of inquiry is the same: deduction of hypotheses (from a theory or a program), test, and subsequent modification (of the theory or the program). Both are attributive, both are value-bound, and both employ the standard "inquiry cycle" (Marx & Hillix, 1963) in pursuit of knowledge. On this "judgmental attribution" issue, social psychology and social evaluation appear to be highly homomorphic.

The second argument contends that there exists a multiplicity of concerned parties in social evaluation settings with vital interests in the design, collection and disposition of data relevant to any program or project. These parties, because of formal or informal role positions, ordinarily possess conflicting values and preferences about any research endeavor, and will attempt influence on the researcher with respect to the focus, design and conduct of the evaluation (Argyris, 1970; Evans, 1972; esp. Krause & Howard, in press; Roston, 1973; Taylor, 1973; Weiss, 1973). Consequently, a grand conflict of research interests may ensue, with the researcher either ineffectively caught in the middle or else enlisted as a partisan for some factional cause. It is claimed that the common existence of such conflicts renders the traditional research model emphasizing dispassionate objectivity and especially total experimenter control incapable of implementation (Guttentag, 1973; Krause & Howard, in press; Weiss, 1973).

Even if we grant the premises, it is not clear that the existence of meta-conflicts in social evaluation settings renders these distinct from traditional social psychological settings. Though few traditionalists admit it, the university as a setting for basic research is also quite well described by the meta-conflict paradigm (cf. Wolfe, 1971). The assistant professor at a large university, like his evaluator counterpart, must also contend with competing research factions, whether from the chairman's discouragement of "unfundable projects," students' reluctance to play captive guinea pig, or just the preferences of a journal editor who regards the investigator's interest area as passé. This is not to say that either traditional social psychological research or social evaluation studies are vacuous proclamations "bought" by the most potent power broker in either setting. Rather, we just recognize the reality that researchers in university labs as

well as those evaluating ongoing social programs must contend with a number of parties who want, and who will exert pressure to obtain, different processes and products from the researcher. Traditional or evaluational, research is always to some degree a negotiated compromise. This distasteful aspect of "doing science" has persisted from the time of the Greek scientists and their patrons to the present. It offers no substantive differentiating criterion by which social evaluation may be segregated from traditional research endeavors (cf. also Taylor, 1973; Weiss, 1973).

The third, political sensitivity, argument asserts that social evaluation researches are always carried out within, and affected by, the political or bureaucratic system in which the target program is embedded. Consequently, social evaluation efforts often fall victim to system sensitivities having no direct relevance to the research project, but which may impede investigatory efforts (e.g., David, 1971; Evans, 1972; Taylor, 1973; Weiss, 1973). For example, research designs potentially yielding information which would reflect unfavorably on program administrators and sponsors may be rejected, or unflattering findings may be suppressed. Even the decision of which program to evaluate is political (Weiss, 1973), since more successful programs are more likely to be subjected to scrutiny because of pressure from public officials needing favorable political ammunition. These "secondary concerns," because of their biasing effects toward administrative protection and research censureship, are said to render the traditional model inapplicable.

Again, however, there appears to be no lack of good homomorphism between the social evaluation arena and traditional social psychological settings. Anyone who has been a nontenured part of an academic psychology department is well aware of the particularly political nature of his continuing appointment (Wolfe, 1971).[4] Structurally, the size and composition (e.g., sex

ratios) of the department, and even the secretarial assistance provided are all political decisions which are "irrelevant to," but impact on, the research endeavor. Functionally, some problem areas are considered highly threatening by administrators (e.g., race or pornography research), and may be discouraged. And, strong legitimation effects exist as well, as when a prominent investigator opens a new area and determines for others a "worthwhile" problem. It is unpleasant and not very tactful to raise these points about traditional research. Unless they are at least broached, however, we are in danger of further fractionating the discipline just to maintain a set of convenient fictions.

The final "cardinal difference" posed is the instability argument. It is said that social evaluation research is performed on programs or projects which are inherently unstable since they are designed to be both adaptive and evolutionary. That is, services and programs are client-centered--they exist to deliver service, and hence innovate, adapt, change, and mutate constantly in order to meet that aim. Thus, the traditional emphasis on constancy of measurable phenomena and the establishment of controllable treatments is simply not applicable to social evaluation pursuits, since these display temporal and system instabilities rendering them immune to the sorts of research forays usually launched by traditionalists (Evans, 1972; Guttentag, 1973; Krause & Howard, in press; Taylor, 1973).

Clearly, one could appeal to the field research tradition in classical social psychology to establish homomorphism between the adaptive systems and investigatory difficulties of evaluational and traditional endeavors (cf., e.g., Webb et al., 1966). However, similarities can be found using very "basic" examples as well. Those readers familiar with the study of choice behavior via some formal econometric model will recognize all the instabilities

claimed as province by social evaluation writers inherent in this "lab" research setting. Concerning structural factors, no one is quite clear what utility or probability functions look like. Regarding adaptation and evolution, human decision makers must be viewed as "functional gain-maximizers," information-processers that learn (change) from experience (e.g., Bonoma, in press) and show a shocking lack of concern for controllability and stability as well. The essential point is just that humans, quite pleasantly, are unstable (i.e., adaptive) systems of the most refined order. Regardless of the point or level of application, any investigatory effort which deals with their behaviors must necessarily and simultaneously encounter system change parameters. These parameters may be more complex and confounded in evaluation than in (say) the study of choice behavior, but are nonetheless clearly "pattern-matched" (i.e., homomorphic) whether it is the adaptive system of a social agency or that of a single human being which is observed.

Experimentation, The Scientific Method, and The Logic of Inquiry

Because of these alleged differences, writers on social evaluation (cf. esp. Guttentag, 1975; Krause & Howard, in press) contend that the usual strategy of inquiry in social psychology, that of the experiment, is unsuitable to evaluation pursuits. As Guttentag describes it, the social evaluation researcher's hypotheses must be translated into an impossible null testing format, preposterous assumptions of randomness made, meaningless tests of significance employed, and the entire effort rendered futile by squeezing it into "a classical experimental straight-jacket" (p. 4). Krause and Howard argue the same point in a more detailed manner. They claim that (a) it is impossible to select a set of independent variables which produce replicable effects in a service program; (b) it is impossible

to demonstrate the construct validity of any set of variable operationalizations; and (c) the complete variable set describing a program must necessarily be unknown.

Contrary to these conclusions, the comparisons suggested above indicate that there quite possibly exists a fair-to-good degree of homomorphism between the nature and context of both traditional social psychological and social evaluation research settings.[5] However, our counter-arguments emphasizing the similarities rather than differences between social evaluation and traditional social psychological research in no way weaken the evaluation researchers' contention that the existence of conflicts, sensitivities and instabilities degrades the applicability of the experimental method to these settings. Rather, they may be correct in this hypothesis, but for the wrong reasons. Since social evaluation settings appear to be homomorphic to traditional concerns, their contention that system conditions often make the experimental method incapable of implementation opens this Pandora's box for all social psychology to a greater or lesser degree. Our comments here, then, are directed toward both settings.

We do not enter this controversy at its most general level, because several recent and excellent pieces exist by both supporters (e.g., Schlenker, 1974) and detractors (Gergen, 1973, in press; cf. also Koch, 1971) of the experimental method and the inquiry process as it is currently practiced in social psychology. If one accepts the initial assumption that social events are at least partially orderly (i.e., causally produced), then it can be demonstrated (cf. Kaplan, 1964) that such phenomena can be most efficiently studied by way of the experiment, given certain intial conditions. The existence of the the appropriate "initial conditions" is exactly what writers such as Gergen (1973), Koch (1971) and Newell (1973) question, and contrasting

their views with the arguments of social evaluation researchers shows the latter arguments to be specific forms of the general questions raised by the former. For our purposes, the question translates as, "Does the existence of meta-conflicts of interest, political sensitivities and a focus on unstable phenomena degrade the applicability of experimental investigatory strategies in traditional as well as evaluation applications?"

Contrary to Schlenker (1974), we believe the answer must be a qualified "yes." That is, and in partial agreement with Gergen and Koch, it is true that the existence of conditions such as those cited by social evaluation specialists often renders experimentation, with its requirement of rigorous control, impractical or impossible. Moreover, this is often the case in social evaluation as well as traditional social psychological research settings. The qualifications, and our disagreement with segregationists (and to some extent, with those who view social psychology as history), are that: (1) these critics often choose to focus on only the simpler, and hence most degradable, experimental methods in their attacks; (2) they adopt a rather parochial view of what constitutes the scientific method generally; and (3) social evaluation authors in particular seem to confuse the scientific method with the logic of scientific inquiry in psychology.

Concerning the first point, it is safe to say that the scientific method (e.g., Kaplan, 1964; Marx & Hillix, 1963) includes more options under the heading "experiment" than the typical garden variety factorial design. Sophisticated and non orthogonal designs, many incorporating intentionally confounded factors and temporal variations, exist which are more suited for investigating complex as opposed to simple phenomena (cf. Campbell & Stanley, 1963; Winer, 1971). Statistical assumptions and data analytic procedures are ordinarily robust enough to allow both ad hoc treatment

assignments and the collection of only the most raw sorts of ordinal measures with little loss of inference ability (cf. Games & Klare, 1968). When even such "relaxed" experiments cannot be attempted due to system constraints, experimental tactics include as full-fledged members certain uses of the simulation (cf. Abelson, 1968) and other operating representations of complex systems. And, when the system approached is only marginally understood in its full complexity, the more advanced correlational (e.g., covariance, path analysis, time-lagged analysis) designs and data partitioning techniques allow teasing out experimental effects from system noise. These experimental techniques are available to our current specialists, both in the traditional and social evaluation areas, and require neither new researchers nor novel methods for their implementation.

It remains true, however, that even the more sophisticated experimental designs may be basically incompatible with the system constraints existing in social evaluation settings. This is because of the nature of the experiment, which is designed to serve as a "snapshot" of effects produced under specified system states. Taking such a picture is of little value if the subject changes immediately after exposure, and of very little value at all if such changes are the result of systematic differences in powerful factors assigned to "error variance" (e.g., political sensitivities) completely outside the realm of experimental interest. Therefore, the usual strategy of inquiry, which includes forming a rudimentary "map" (i.e., theory) of the investigatory area and then exploring this map via experimentation, may be degraded for the reason that the entire map is nonrandomly affected by its enclosure within progressively larger systems.

It is not the existence of systems within systems per se which degrades the application of experimentation; as Shooster (1971) has pointed out, even

extremely complex systems are amenable to classification and experimentation. Rather, the basic problem is to specify an appropriate level of analysis for "map" formation before experiments are undertaken: it makes little sense to attempt theory-formation about (say) small group interactions within a service or program when these are known to be nonrandomly affected by the sort of macrosystem constraints addressed by social evaluation methodologists (cf. also Laszlo et al., 1974). Unfortunately, detailed knowledge of system confounds and constraints is ordinarily not possessed for any specific setting prior to investigation--such knowledge only accrues through repeated unsuccessful investigatory attempts. Social evaluation researchers, then, directly encounter the "dilemma of complex systems" in their avocation: a sufficiently articulated theoretical "map" of the investigatory area is needed to permit highly controlled observations, but good experimental observations must await a detailed map which distinguishes exprimental effect from system noise. This dilemma may be presumed to have led both to social evaluation's alienation from the experimental method and from traditional social psychology.

If the current arguments regarding the homomorphism of evaluation and traditional social psychology are valid, we should expect that traditional endeavors encounter a similar form of the dilemma of complex systems. This appears to be the case. With regard to dissonance phenomena, for example, it remains (after 17 years) impossible to specify complete variable sets or even highly probable functional relations in current explanatory efforts (cf. Kelman, 1974; Krause, 1972), a phenomenon which parallels the Krause-Howard complaints about evaluation endeavors. Rather retarded progress has been largely observable not only for dissonance but for most traditional social psychology (cf., e.g., Koch, 1971) primarily because the dyad and the

small group have only recently been taken seriously as (1) social systems which (2) are significantly affected by the other systems in which they are embedded. For the first, we now talk about experimenter-subject interactions, when we very recently used to conceive of our discipline as the study of individual responses to social stimuli (Shaw & Costanzo, 1970: p. 3). With regard to the second, we currently seek knowledge which goes beyond the college sophomore population in its generality, realizing that the various macrosystems in which subjects and observations are embedded is an integral part of "objective" data production (cf. Miller et al., in press). This change in thinking has been no less than revolutionary for traditional social psychology, and has led to the dual result of producing more generalizable data while simultaneously awakening traditionalists to the awareness that they must confront system dilemmas as well. The dilemma of complex systems has been less visible in traditional social psychology just because such research attempts to minimize cross-system influences (i.e., to maximize control) in maximally simple social systems (e.g., dyads). That such minimization could not (and should not have been expected to) eliminate all system influences is a phenomenon with which we are just now learning to deal (e.g., Kelman, 1974).

Thus, simple experimentation in both the social evaluation and traditional social psychological domains may be degraded by system constraints impinging on the investigatory area. This "dilemma of complex systems" is often more poignantly experienced by social evaluation researchers than by traditional social psychologists because (1) the former do not (and cannot) attempt to minimize cross-system influences, and (2) the latter ordinarily restrict themselves to the least complex, and hence most controllable, systems. In fairness to social evaluation authors, we must agree that greater

degradation of the experimental method often is experienced in program evaluation than in the study of dyadic behavior. However, the low level and power of traditional social psychological laws (cf. Gergen, 1973) attests at least partially to the existence of a homomorphic dilemma within the traditional domain as well. Thus, the settings and problems of the one cannot logically be segregated from those of the other on these grounds. It can be concluded that no investigatory-explanatory approach which ignores the nexus of systems in which the phenomenon of interest is embedded can produce generalizable knowledge. Further, it may be expected that both traditional and evaluational experimental pursuits will gain in feasibility with (1) the methodological sophistication of the researcher, and (2) the articulation of general principles of cross-system influence (cf., e.g., Grinker, 1967).

While a systems theory approach to interactive events may eventually offer better "maps" which will increase experimental applicability and results generalizability, other tactics subsumed in the scientific method offer more immediate application. Most prominent among the usually ignored observation techniques detailed in every introductory text is that of naturalistic observation. As Charlesworth (1973) points out, the difference between experimental studies and naturalistic observation is just that the former concentrate on what an organism (or organization) can do under specified and known system conditions, while the latter concentrates on what an organism does do under the operative system constraints in its environment (even if these are currently unknown). Naturalistic observation, additionally offers advantages beyond a simple method of observation. The work of a number of serious students of social ethology (e.g., Bales, 1971; Barker, 1963; Charlesworth, 1973) suggests that the raw behavior of humans

in their social habitats often divides itself into theoretically useful categories and typologies (e.g., Barker's behavior settings). Therefore, such studies could serve the dual purpose in both social evaluation and traditional social psychology not only of establishing a raw data base of ongoing behavior, but also of providing strong pre-experimental indicators of which system facets and behavioral regularities are of crucial theoretical concern.

Again, the case history method, while ordinarily not seriously considered by traditionalists, may serve as an extremely valuable tool in both traditional and evaluation studies by which to circumvent the dilemma of complex systems. The collection of a wide variety of historical and current reports by members of a service unit (or parties to social interaction) may bear the imprinteur of secondary data as opposed to more "behavioral" measures. Case studies have, however, the advantage of permitting economical comparisons between the numerous competing factions encountered in complex systems, as well as quickly pointing out "deviant cases" among factions for special theoretical or subsequent experimental focus. When combined with naturalistic observation of a service unit, such techniques allow at least rough convergent validity estimates to be performed on otherwise unmeasurable system or behavioral components (cf. Rapoport, 1968).

Thus, the second major point is this: writers on social evaluation, as well as some of those critical of traditional social psychology, have taken a rather parochial viewpoint by equating certain restricted aspects of the experimental method with the scientific method. Then, because it can be argued that application of the former is often degraded by conditions found to exist in complex systems, it is concluded that novel methodological techniques and new specialists are demanded. However, other "standard" investigatory techniques in science are available which are not as prone

to system degradation as the experiment, for they do not require the same stringent control specifications and are more suited to exploration in the absence of a well-articulated theoretical "map." The segregatory conclusion of the more extreme social evaluation authors is unwarranted: rather, it is only necessary to look beyond the n-group factorial paradigms on which we are functionally fixated to the remainder of the techniques comprising the scientific method. Those who criticize transplantation of traditional simple experimentation to complex evaluation settings are correct when they assert that these are ordinarily degraded there, but incorrect when they conclude that new methods are urgently needed or a different subdiscipline thereby comes into existence. Conversely, those who always choose their basic research problems to fit the already-prepared bed of orthogonal variance estimates are equally incorrect, since they fail to perceive that traditional social psychological research is in principle no less a social influence setting, no less political, no less unstable than a confounded service agency.

Concerning the final point, it appears that social evaluation writers have chosen to equate the scientific method with what is ordinarily known as the strategy of inquiry in science (e.g., Homans, 1967; Marx & Hillix, 1963). That is, the inquiry process is usually thought of as having two stages: discovery and explanation. Those tactics included in the scientific method (experimentation, naturalistic observation, etc.) facilitate the discovery of "facts" in a manner (e.g., replicability, public verifiability) consistent with scientific endeavors, while the mechanisms of law, model and theory allow subsuming such facts under more general explanatory principles. It is these principles, when convergently confirmed by different operationalizations and across system levels, which are expected to be generalizable

to theoretically similar phenomena. To the extent to which social evaluation researchers have attempted to apply experimentation directly to public sector phenomena, neglecting the prior steps of hypothesis derivation and the succeeding steps of theory revision, they short-cut the scientific method. And, whenever the logic of inquiry is short-cut either by design or system constraints, nongeneralizable knowledge is the result. Experimental results per se were never intended to be generalizable to even moderately similar phenomena (cf., e.g., Marx & Hillix, 1963). Rather, these discovery tools provide one of a number of ways to accrue facts in the service of theoretical (i.e., explanatory) notions. These theoretical propositions, and only these, are expected to be generalizable to between-system phenomena other than the one currently being investigated. At base, then, those who would claim the scientific method is inapplicable to complex social settings are (justifiably) lamenting the lack of theoretical principles to guide them in their study, and confirming first-hand that the discovery tools of science cannot be applied in the absence of the explanatory.

In sum, the present view argues that the settings and problems of traditional and social evaluation research are homomorphic, and consequently have similar (but not completely equivalent) consequences regarding the applicability of experimentation in particular and the scientific method in general. Experimentation often cannot be easily applied to social evaluation problems because of the "dilemma of complex systems:" however, a less-recognized but homomorphic dilemma exists for traditional research as well. Fortunately, other standard discovery methods exist less subject to system degradation. Regardless of the discovery technique applied, though, collected data must be incorporated within viable theory in order to be generalizable.

Inquiry and Relevance

Articulation of the homomorphic but differentially intense dilemma of complex systems as it affects traditional social psychological studies and social evaluation pursuits offers a clarifying and integrative set of suggestions regarding the path of research in each setting. When combined with the stereotypic motivational biases of traditionalists versus social evaluation researchers (e.g., Evans, 1972; Meehl, 1972), it offers as well some commentary regarding the achievement of "relevance" in either endeavor. We turn to such considerations by way of conclusion.

In the present view, arguments about the segregation of social evaluation from traditional pursuits, alleged motivational deficits in one or another camp, and differential claims (and disclaimers) of relevance all stem from an improper understanding of the strategy of scientific inquiry as it applies to human social behavior. Put in oversimple fashion, researchers employ the scientific method in order to add support or to aid in modifying theories of social behavior. No one research application is relevant in the sense of being directly applicable to the real world, but a "net" of empirically confirmed and extended theoretical propositions may be subsequently and rightly employed to generate a coherent set of decision rules. These rules, if generated from established and cross-validated theory, will carry enough external validity to focus observational strategy, suggest variable sets, and so on when similar but more complex (i.e., cross-system) phenomena are addressed.

However, the construction of sound theory in either setting may be approached in somewhat different fashion due to differences in the intensity of the dilemma of complex systems. Social evaluation researchers, which deal with maximally complex phenomena embedded in recursive and reflexive

systems, will probably initially find naturalistic observation and the case method initially less degradable than experimentation per se. Traditional social psychology, with its focus on simple systems ordinarily taken "out of context," can easily continue to apply experimentation techniques as long as there is an increasing recognition that system interactions cannot be "controlled out" of observation settings. Thus, social evaluation and social psychology may initiate research projects in somewhat opposite manners for "best" theoretical progress: the former might start with naturalistic observations and case studies to delineate the rough system boundaries, and progress to experimentation only as cross-system influences become more clear. The latter might initiate rigorous experimentation as the most efficient method of answering causal questions in controlled microsystems, but must relax its methods periodically and convergently to determine if the questions asked have any relation to those occurring in the nexus of systems in which the phenomenon is ordinarily embedded.

It is clear that both traditional social psychology and social evaluation research are impaled on different horns of the same dilemma of complex systems. That their problem settings force opposite sides and different manifestations of the dilemma should not blind either to the fact that (1) this in no substantive way distinguishes the efforts of the one from the other; (2) the "nonpreferred" side of the dilemma must eventually be addressed; and (3) no artificial segregation of evaluation from traditional pursuits can lead to the successful resolution of the dilemma of complex systems. Whether one's initial preferences lie with articulated maps of systems so simple and controlled that the resultant rigorous knowledge is not very generalizable, or with complex real world endeavors yielding only the most confounded indices in the absence of a good "map," only the convergent

pursuit of both approaches can ever result in the understanding and prediction of interactive events. It follows conclusively that to segregate social evaluation from social psychology is not only illogical because of their high degree of homomorphism, but eventually will be fatal for both as well.

Regarding differential claims to relevance, it is clear that relevance resides not in the observational field, but in the resultant theory. Neither traditionalists nor social evaluation researchers have reason to denigrate the activities of the other, since we have seen both are connected by a common problem. To the extent to which social evaluation researchers are stereotypically represented as more interested in immediate applications of knowledge than traditional social psychologists, these have historically felt justified to claiming "face" relevance (i.e., first-order relevance: Meehl, 1972). Conversely, traditionalists, because they currently possess the most viable of our theories in social psychology, have felt justified in claiming "construct" relevance (i.e., second-order relevance).[6] Our present recognition that both encounter the dilemma of complex systems in somewhat different form argues strongly that neither faction should be too eager to disclaim the other's motivations, since both will eventually encounter the problems of the other. Further, and if performance is to judge, neither faction has been very relevant at all. The real world applications of evaluation specialists have more often resulted in confusion than clarity about public sector programs, while the theories of traditionalists have often been so system-bound as to be ungeneralizable. Since both "factions" appear to have the same ultimate motivation, albeit through different implementation strategies, it would seem that differential claims to relevance are unfounded.

More important, and partially connected, is that history will judge only those of our number as relevant who participate in the development of articulated theory about interactive events. Articulated theory is not achieved by making a choice about whether to focus on microsystems for precise map formation, or on real world programs to garner generalizable data. Rather, only when the dilemma of complex systems is resolved in the accomplishment of both these facets in one social psychology will differential relevance claims become feasible. These will then consist of pitting competing theories against one another, not competing researchers.

References

Abelson, R. P. Simulation of social behavior. In G. Lindzey & E. Aronson (Eds.), The handbook of social psychology. 2nd Edition. Vol. II. Reading, Mass.: Addison-Wesley, 1968. Pp. 274-356.

Argyris, C. Intervention theory and method: A behavioral science view. Reading, Mass.: Addison-Wesley, 1970.

Bales, R. F. Personality and interpersonal behavior. New York: Holt, 1971.

Barber, T. X. and Silver, M. J. Pitfalls in data analysis and interpretation: A reply to Rosenthal. Psychological Bulletin Monograph Supplements, Part 2, December, 1968. Pp. 48-62.

Barker, R. G. The stream of behavior. New York: Appleton-Century-Crofts, 1963.

Bonoma, T. V. A new methodology for the study of individual and social choice behavior. Journal for the Theory of Social Behaviour, in press.

Campbell, D. T. and Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand-McNally, 1963.

Charlesworth, W. R. Ethology's contribution to a framework for relevant research. Paper presented at the 81st annual meeting of the American Psychological Association, Montreal, August 27-31, 1973.

Cook, D. L. Program evaluation and review technique (PERT). Office of Education Cooperative Monograph No. 17, 1966.

David, E. E., Jr. Making objectivity credible and acceptable. Invited address presented at the 79th annual meeting of the American Psychological Association, Washington, D.C., September, 1971.

DuBois, P. H. and Mayo, G. D. (Eds.) Research strategies for evaluating training. AERA Monograph Series on Curriculum Evaluation. Vol. 4. [illegible]: Rand McNally, 1970.

Evans, J. W. Evaluating social action programs. In G. Zaltman, P. Kotler and I. Kaufman (Eds.), Creating social change. New York: Holt, 1972. Pp. 629-639.

Games, P. A. and Klare, G. R. Elementary statistics: Data analysis for the behavioral sciences. New York: McGraw-Hill, 1968.

Gergen, K. Social psychology as history. Journal of Personality and Social Psychology, 1973, 26, 309-320.

Gergen, K. J. Social psychology, science and history: A rejoinder. Journal of Personality and Social Psychology, in press.

Grinker, R. R., Sr. (Ed.) Toward a unified theory of human behavior. New York: Basic, 1967.

Grobman, H. Evaluation activities of curriculum projects. AERA Monograph Series on Curriculum Evaluation. Vol. 2. New York: Rand-McNally, 1970.

Guttentag, M. Subjectivity and its use in evaluation research. Paper presented at the annual meeting of the Society for Psychotherapy Research, Philadelphia, June, 1973.

Homans, G. C. The nature of social science. New York: Harcourt, 1967.

Isaacson, R. A., Bunker, B. B., Burke, W. W., Gindes, M. and Lewicki, R. F. Social intervention: A behavioral science approach. New York: Free Press, 1971.

Hull, C. The conflicting psychologies of learning--a way out. Psychological Review, 1935, 42, 491-516.

Kaplan, A. The conduct of inquiry. San Francisco: Chandler, 1964.

Kelman, H. C. Attitudes are alive and well and gainfully employed in the sphere of action. American Psychologist, 1974, 29, 310-324.

Kemeny, J. G. A philosopher looks at science. Princeton: Van Nostrand, 1959.

Koch, S. Reflections on the state of psychology. Social Research, 1971, 38, 669-709.

Koen, F. M. The evaluation of training programs for college teachers. Paper presented at the 81st Annual Meetings of the American Psychological Association, Montreal, August 27-31, 1973.

Krause, M. S. An analysis of Festinger's cognitive dissonance theory. Philosophy of Science, 1972, 39, 32-50.

Krause, M. S. and Howard, K. I. Program evaluation in the public interest: A new research methodology. Community Mental Health, in press.

Laszlo, C. A., Levine, M. D. and Milsum, J. H. A general systems framework for social systems. Behavioral Science, 1974, 19, 79-82.

Lewin, K. Field theory in social science. New York: Harper & Row, 1951.

Marx, M. H. and Hillix, W. A. Systems and theories in psychology. New York: McGraw-Hill, 1963.

Meehl, P. E. Second-order relevance. American Psychologist, 1972, 27, 932-940.

Miller, R. R., Brickman, P. and Bolen, D. Attribution versus persuasion as a means of modifying behavior. Journal of Personality and Social Psychology, in press.

Moores, D. F. Moving research across the relevancy continuum. Paper presented at the 81st Annual Meetings of the American Psychological Association, Montreal, August 27-31, 1973.

Nagel, E. The structure of science. New York: Harcourt, 1961.

Nelson, B. Psychologists: Searching for social relevance at APA meeting. In. R. E. Buckhout et al. (Eds.), Toward social change. New York: Harper & Row, 1971. Pp. 60-62.

Newell, A. You can't play 20 questions with nature and win. In W. G. Chase (Ed.), Visual information processing. New York: Academic Press, 1973. Pp. 283-310.

Proshansky, H. M. The environmental crisis in human dignity. Journal of Social Issues, 1974, 29, 1-20.

Rapoport, A. Prospects for experimental games. Journal of Conflict Resolution, 1968, 12, 461-470.

Roston, R. A. Moral certitude and "technical validities." Paper presented at the 81st Annual Meetings of the American Psychological Association, Montreal, August 27-31, 1973.

Samuels, S. J. How psychologists can be relevant or you can have your cake and eat it too. Paper presented at the 81st Annual Meetings of the American Psychological Association, Montreal, August 27-31, 1973.

Schlenker, B. R. Social psychology and science. Journal of Personality and Social Psychology, 1974, 29, 1-15.

Scriven, M. The methodology of evaluation. In R. Tyler, R. Gagne, & M. Scriven (Eds.), Perspectives of curriculum evaluation. AERA Monograph Series on Curriculum Evaluation. Vol. 1. New York: Rand McNally, 1967. Pp. 39-83.

Sechrest, L. Training in evaluation research: The development of a program. Paper presented at the 81st Annual Meetings of the American Psychological Association, Montreal, August 27-31, 1973.

Shaw, M. E. and Costanzo, P. R. Theories of social psychology. New York: McGraw-Hill, 1970.

Shooster, C. N. Tests and prediction: A systems analysis approach. Behavioral Science, 1974, 19, 111-118.

Silverman, I. Crisis in social psychology: The relevance of relevance. American Psychologist, 1971, 26, 583-584.

Taylor, J. B. Is evaluation sufficient: A possible future. Paper presented at the 81st Annual Meetings of the American Psychological Association, Montreal, August 27-31, 1973.

Webb, E. J., Campbell, D. T., Schwartz, R. D. and Sechrest, L. Unobtrusive measures: Nonreactive research in the social sciences. Chicago: Rand-McNally, 1966.

Weiss, C. H. Evaluation research. Englewood Cliffs, N. J.: Prentice-Hall, 1972.

Weiss, C. H. Evaluation research in the political context. Paper presented at the 81st Annual Meetings of the American Psychological Association, Montreal, August 27-31, 1973.

Winer, B. J. Statistical principles in experimental design. (2nd ed.), New York: McGraw-Hill, 1971.

Wolfe, A. The myth of the free scholar. In R. B. Buckhout et al. (Eds.), Toward social change. New York: Harper & Row, 1971. Pp. 64-67.

Wozniak, R. H. In-context research on children's learning as a basic science prophylactic: Or true purity doesn't need to wash. Paper presented at the 81st Annual Meetings of the American Psychological Association, Montreal, August 27-31, 1973.

Footnote

1. Requests for reprints may be sent to the author at the Institute for Juvenile Research, Department of Mental Health, 1140 South Paulina Street, Chicago, Illinois, 60612. The author wishes to express his gratitude to Merton Krause and Kenneth Howard for discussing their work, examining his, and encouraging the differences. This paper is not intended to represent the policy or views of the Illinois Department of Mental Health with regard to social evaluation research.

Footnotes

1. (on separate page).

2. We restrict our present comments primarily to the more recent (1972-1974) treatments of social evaluation research. Much earlier literature exists, especially from education and government (e.g., Cook, 1966; DuBois & Mayo, 1970; Grobman, 1970; Scriven, 1967), which treats the evaluation enterprise as a direct extension of traditional research procedures. It is thus at least partially immune to the criticisms raised here. The curious reader might compare an early with a more recent piece on social evaluation to determine for himself the radical changes which have occurred.

3. Traditional here has as its meaning the design and conduct of controlled experimental investigations, usually within a laboratory context.

4. Wolfe (1971) has devastatingly described the conduct of traditional research in a university environment, and the numerous cross-cutting pressures impinging on the so-called "impartial" scholar. He concludes, "We have now arrived at a definition of a successful scholar. He is a person who constantly reiterates different aspects of the same idea in a manner determined for him by others without being critical of the conditions which shaped his life" (p. 66). While it is true that evaluation specialists are "selling" a research product, it is no less true of the academician (e.g., the emphasis on statistically "significant" results).

5. It is, however, quite correct to argue that either the greater frequency of such problems, or the presence of a greater number of such issues, renders the social evaluation setting practically if not principally distinct from the traditional (see below).

6. Though one should not single out Meehl here, since he just nicely stated a distinction recurrent in the literature since the '40's (e.g., Marx & Hillix,

1963), the artificial positing of different "sorts" of relevance is both an indicator of a basic misunderstanding of the inquiry process and an early sign of the segregatory disease.